SEQUENCE LISTING

<110> WALLAART, Thorvald Eelco
BOUWMEESTER, Hendrik Jan

<120> Transgenic Amorpha-4, 11-Diene Synthesis

<130> 702 010272

<140> 09/763,822
<141> 2001-02-26

<150> PCT/EP99/06302
<151> 1999-08-27

<160> 14

<170> MS Word 97 SR-2

<210> 1
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> EcoR I (Not I) adapter

<400> 1
gtcgacgcgg ccgcg 15

<210> 2
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> EcoR I (Not I) adapter

<400> 2
cagctgcgcc ggcgcttaa 19

<210> 3
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Sense primer (primer C) used in PCR amplification

<400> 3
gtcgacaaac catggcactt acagaag 27

<210> 4
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Antisense primer (primer D) used in PCR amplification

<400> 4

ggatggatcc tcatatactc ataggataaa cg 32

<210> 5
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Sense primer (primer G) used in PCR amplification

<400> 5
gaggatccat gtcacttaca gaa 23

<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Antisense primer (primer H) used in PCR amplification

<400> 6
atggatcctc atatactcat agga 24

<210> 7
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Sense primer (primer E) used in PCR amplification

<400> 7
cgagaattca tgtcacttac ag 22

<210> 8
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Antisense primer (primer F) used in PCR amplification

<400> 8
ggatctcgag tcatatactc at 22

<210> 9
<211> 538
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of probe generated by PCR with primers A and B

<400> 9
gatgagaatg ggaaatttaa ggaatcgtta gctaatgatg ttgaaggttt gcttgagttg 60

tacgaagcaa cttctatgag ggtacctggg gagattatat tagaagatgc tcttggtttt 120

acacgatctc gtcttagcat tatgacaaaa gatgcttttt ctacaaaccc cgctcttttt 180

accgaaatac aacgggcact aaagcaaccc ctttggaaaa ggttgccaag aatagaggcg 240

gcgcagtaca ttcctttcta tcaacaacaa gattctcata acaagacttt acttaaactt 300

gctaagttag agttcaattt gcttcagtca ttgcacaagg aagagctcag ccatgtgtgc 360

aaatggtgga aagctttcga tatcaagaag aacgcacctt gtttaagaga tagaattgtt 420

gaatgctact tttggggact aggttcaggc tatgagccac agtattcccg ggctagagtt 480

ttcttcacaa aagctgttgc tgttataact cttatagacg acaccttcga cgctacgg 538

<210> 10
<211> 179
<212> PRT
<213> Artificial Sequence

<220>
<223> Deduced amino acid sequence of probe generated by PCR with primers A and B

<400> 10

Asp	Glu	Asn	Gly	Lys	Phe	Lys	Glu	Ser	Leu	Ala	Asn	Asp	Val	Glu	Gly
1				5				10					15		
Leu	Leu	Glu	Leu	Tyr	Glu	Ala	Thr	Ser	Met	Arg	Val	Pro	Gly	Glu	Ile
		20					25					30			
Ile	Leu	Glu	Asp	Ala	Leu	Gly	Phe	Thr	Arg	Ser	Arg	Leu	Ser	Ile	Met
	35					40					45				
Thr	Lys	Asp	Ala	Phe	Ser	Thr	Asn	Pro	Ala	Leu	Phe	Thr	Glu	Ile	Gln
50					55					60					
Arg	Ala	Leu	Lys	Gln	Pro	Leu	Trp	Lys	Arg	Leu	Pro	Arg	Ile	Glu	Ala
65				70					75						80
Ala	Gln	Tyr	Ile	Pro	Phe	Tyr	Gln	Gln	Gln	Asp	Ser	His	Asn	Lys	Thr
			85					90					95		
Leu	Leu	Lys	Leu	Ala	Lys	Leu	Glu	Phe	Asn	Leu	Leu	Gln	Ser	Leu	His
		100					105					110			
Lys	Glu	Glu	Leu	Ser	His	Val	Cys	Lys	Trp	Trp	Lys	Ala	Phe	Asp	Ile
	115					120					125				
Lys	Lys	Asn	Ala	Pro	Cys	Leu	Arg	Asp	Arg	Ile	Val	Glu	Cys	Tyr	Phe
130					135					140					
Trp	Gly	Leu	Gly	Ser	Gly	Tyr	Glu	Pro	Gln	Tyr	Ser	Arg	Ala	Arg	Val
145				150					155						160
Phe	Phe	Thr	Lys	Ala	Val	Ala	Val	Ile	Thr	Leu	Ile	Asp	Asp	Thr	Phe
			165					170					175		
Asp	Ala	Thr													

<210> 11
<211> 2112
<212> DNA
<213> Artemisia annua L.

<220>
<223> Nucleotide sequence of a positive clone (amorphadiene synthase encoding gene) isolated from the cDNA library of induced A.annua

<400> 11
aattcgcggc cgcgtcgaca aatcatgtca cttacagaag aaaaacctat tcgccccatt 60

gccaactttc ctccaagcat ttggggagat cagtttctca tctatcaaaa gcaagtagag 120

caaggggtgg aacagatagt gaatgattta aaaaaagaag tgcggcaact actaaaagaa 180

gctttggata ttcctatgaa acatgccaat ttgttgaagc tgattgatga aattcaacgc 240

cttggaatac cgtatcactt tgaacgggag attgatcatg cattgcaatg tatttatgaa 300

acatatggtg ataactggaa tggtgaccgc tcttccttat ggttccgtct tatgcgaaag 360

caaggatatt atgttacatg tgatgttttc aataactata aagacaaaaa tggagcgttc 420

aagcaatcgt tagctaatga tgttgaaggt ttgcttgagt tgtacgaagc aacttctatg 480

agggtacctg gggagattat attagaagat gctcttggtt ttacacgatc tcgtcttagc 540

attatgacaa aagatgcttt ttctacaaac cccgctcttt ttaccgaaat acaacgggca 600

ctaaagcaac ccctttggaa aaggttgcca agaatagagg cggcgcagta cattcctttc 660

tatcaacaac aagattctca taacaagact ttacttaaac ttgctaagtt agagttcaat 720

ttgcttcagt cattgcacaa ggaagagctc agccatgtgt gcaaatggtg gaaagctttc 780

gatatcaaga agaacgcacc ttgtttaaga gatagaattg ttgaatgcta cttttgggga 840

ctaggttcag gctatgagcc acagtattcc cgggctagag ttttcttcac aaaagctgtt 900

gctgttataa ctcttataga tgacacttat gatgcgtatg gtacttatga agaacttaag 960

atctttactg aagctgttga aaggtggtca attacatgct tagacacact tccagaatac 1020

atgaaaccga tatacaaatt attcatggat acatacacag aaatggaaga atttcttgca 1080

aaggagggaa gaacagatct atttaactgc ggcaaagaat ttgtgaaaga gtttgttaga 1140

aacctgatgg ttgaagcaaa atgggcaaat gagggacaca taccaaccac tgaagagcat 1200

gatccagttg taatcattac tggcggtgct aacctgctta caacaacttg ttatcttggc 1260

atgagtgata tattcacaaa agagtctgtc gaatgggctg tctctgcacc tcctcttttt 1320

agatactcag gtatacttgg tcgacgccta aatgatctca tgacccacaa ggccgagcaa 1380

gaaagaaaac atagttcatc gagccttgaa agttatatga aggaatataa tgtcaatgag 1440

gagtatgccc aaaccttgat ttacaaggaa gtagaagatg tgtggaaaga tataaaccga 1500

gagtacctca caactaaaaa cattccaagg ccgttattga tggctgtgat ctatttgtgc 1560

cagtttcttg aagttcaata tgcaggaaag gataacttca cacgtatggg agacgaatac 1620

aaacatctca taaagtctct actcgtttat cctatgagta tatgactacc aatccttcgt 1680

gcatagccta tcaattatat tgaaagggtt aactatgcac gtctctatgg agagaatttc 1740

tcaagctatt tggtgtttct tgctggcaat aataaatcag acgcataaaa ttgtattgaa 1800

ctatatgccg atagctattt aaagttatta tacaactaaa atattcaaca atggtattat 1860

acttttactt tgtacaaaag caaaagtaca ctactgttat gtaacatttt agttctatga 1920

tactttagtt acgaatcggc ttatatacat tgatacactt ttatgcagaa aaccctagta 1980

aataaaaagt cgatatcttg tactacacat atcgcacgaa tttccgtttg ccgtttgtat 2040

tttacgatat gttatttaat gaatatgttt catgtggttg ttgcttaaaa aaaaagtcga 2100

cgcggccgcg aa 2112

<210> 12
<211> 697
<212> PRT
<213> Artemisia annua L.

<220>
<223> Deduced amino acid sequence of a positive clone (amorphadiene synthase encoding gene) isolated from the cDNA library of induced A.annua

<400> 12

```
Asn Ser Arg Pro Arg Arg Gln Ile Met Ser Leu Thr Glu Glu Lys Pro
  1               5                  10                  15
Ile Arg Pro Ile Ala Asn Phe Pro Pro Ser Ile Trp Gly Asp Gln Phe
             20                  25                  30
Leu Ile Tyr Gln Lys Gln Val Glu Gln Gly Val Glu Gln Ile Val Asn
         35                  40                  45
Asp Leu Lys Lys Glu Val Arg Gln Leu Leu Lys Glu Ala Leu Asp Ile
     50                  55                  60
Pro Met Lys His Ala Asn Leu Leu Lys Leu Ile Asp Glu Ile Gln Arg
 65                  70                  75                  80
Leu Gly Ile Pro Tyr His Phe Glu Arg Glu Ile Asp His Ala Leu Gln
                 85                  90                  95
Cys Ile Tyr Glu Thr Tyr Gly Asp Asn Trp Asn Gly Asp Arg Ser Ser
            100                 105                 110
Leu Trp Phe Arg Leu Met Arg Lys Gln Gly Tyr Tyr Val Thr Cys Asp
        115                 120                 125
Val Phe Asn Asn Tyr Lys Asp Lys Asn Gly Ala Phe Lys Gln Ser Leu
    130                 135                 140
Ala Asn Asp Val Glu Gly Leu Leu Glu Leu Tyr Glu Ala Thr Ser Met
145                 150                 155                 160
Arg Val Pro Gly Glu Ile Ile Leu Glu Asp Ala Leu Gly Phe Thr Arg
                165                 170                 175
Ser Arg Leu Ser Ile Met Thr Lys Asp Ala Phe Ser Thr Asn Pro Ala
            180                 185                 190
Leu Phe Thr Glu Ile Gln Arg Ala Leu Lys Gln Pro Leu Trp Lys Arg
        195                 200                 205
Leu Pro Arg Ile Glu Ala Ala Gln Tyr Ile Pro Phe Tyr Gln Gln Gln
    210                 215                 220
Asp Ser His Asn Lys Thr Leu Leu Lys Leu Ala Lys Leu Glu Phe Asn
225                 230                 235                 240
Leu Leu Gln Ser Leu His Lys Glu Glu Leu Ser His Val Cys Lys Trp
                245                 250                 255
Trp Lys Ala Phe Asp Ile Lys Lys Asn Ala Pro Cys Leu Arg Asp Arg
            260                 265                 270
Ile Val Glu Cys Tyr Phe Trp Gly Leu Gly Ser Gly Tyr Glu Pro Gln
        275                 280                 285
Tyr Ser Arg Ala Arg Val Phe Phe Thr Lys Ala Val Ala Val Ile Thr
    290                 295                 300
Leu Ile Asp Asp Thr Tyr Asp Ala Tyr Gly Thr Tyr Glu Glu Leu Lys
305                 310                 315                 320
Ile Phe Thr Glu Ala Val Glu Arg Trp Ser Ile Thr Cys Leu Asp Thr
                325                 330                 335
Leu Pro Glu Tyr Met Lys Pro Ile Tyr Lys Leu Phe Met Asp Thr Tyr
```

```
            340                 345                 350
Thr Glu Met Glu Glu Phe Leu Ala Lys Glu Gly Arg Thr Asp Leu Phe
        355                 360                 365
Asn Cys Gly Lys Glu Phe Val Lys Glu Phe Val Arg Asn Leu Met Val
    370                 375                 380
Glu Ala Lys Trp Ala Asn Glu Gly His Ile Pro Thr Thr Glu Glu His
385                 390                 395                 400
Asp Pro Val Val Ile Ile Thr Gly Gly Ala Asn Leu Leu Thr Thr Thr
                405                 410                 415
Cys Tyr Leu Gly Met Ser Asp Ile Phe Thr Lys Glu Ser Val Glu Trp
            420                 425                 430
Ala Val Ser Ala Pro Pro Leu Phe Arg Tyr Ser Gly Ile Leu Gly Arg
        435                 440                 445
Arg Leu Asn Asp Leu Met Thr His Lys Ala Glu Gln Glu Arg Lys His
    450                 455                 460
Ser Ser Ser Ser Leu Glu Ser Tyr Met Lys Glu Tyr Asn Val Asn Glu
465                 470                 475                 480
Glu Tyr Ala Gln Thr Leu Ile Tyr Lys Glu Val Glu Asp Val Trp Lys
                485                 490                 495
Asp Ile Asn Arg Glu Tyr Leu Thr Thr Lys Asn Ile Pro Arg Pro Leu
            500                 505                 510
Leu Met Ala Val Ile Tyr Leu Cys Gln Phe Leu Glu Val Gln Tyr Ala
        515                 520                 525
Gly Lys Asp Asn Phe Thr Arg Met Gly Asp Glu Tyr Lys His Leu Ile
    530                 535                 540
Lys Ser Leu Leu Val Tyr Pro Met Ser Ile Leu Pro Ile Leu Arg Ala
545                 550                 555                 560
Pro Ile Asn Tyr Ile Glu Arg Val Asn Tyr Ala Arg Leu Tyr Gly Glu
                565                 570                 575
Asn Phe Ser Ser Tyr Leu Val Phe Leu Ala Gly Asn Asn Lys Ser Asp
            580                 585                 590
Ala Asn Cys Ile Glu Leu Tyr Ala Asp Ser Tyr Leu Lys Leu Leu Tyr
        595                 600                 605
Asn Asn Ile Gln Gln Trp Tyr Tyr Thr Phe Thr Leu Tyr Lys Ser Lys
    610                 615                 620
Ser Thr Leu Leu Leu Cys Asn Ile Leu Val Leu Tyr Phe Ser Tyr Glu
625                 630                 635                 640
Ser Ala Tyr Ile His Tyr Thr Phe Met Gln Lys Thr Leu Val Asn Lys
                645                 650                 655
Lys Ser Ile Ser Cys Thr Thr His Ile Ala Arg Ile Ser Val Cys Arg
            660                 665                 670
Leu Tyr Phe Thr Ile Cys Tyr Leu Met Asn Met Phe His Val Val Val
        675                 680                 685
Ala Lys Lys Ser Arg Arg Gly Arg Glu
    690                 695
```

<210> 13
<211> 1649
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the amorphadiene synthase encoding gene, between start and stop codon, obtained by PCR with primers C and D

<400> 13
ccatggcact tacagaagaa aaacctattc gccccattgc caactttcct ccaagcattt 60

ggggagatca gtttctcatc tatcaaaagc aagtagagca aggggtggaa cagatagtga 120

atgatttaaa aaaagaagtg cggcaactac taaaagaagc tttggatatt cctatgaaac 180

atgccaattt gttgaagctg attgatgaaa ttcaacgcct tggaataccg tatcactttg 240

aacgggagat tgatcatgca ttgcaatgta tttatgaaac atatggtgat aactggaatg 300

gtgaccgctc ttccttatgg ttccgtctta tgcgaaagca aggatattat gttacatgtg 360

atgttttcaa taactataaa gacaaaaatg gagcgttcaa gcaatcgtta gctaatgatg 420

ttgaaggttt gcttgagttg tacgaagcaa cttctatgag ggtacctggg gagattatat 480

tagaagatgc tcttggtttt acacgatctc gtcttagcat tatgacaaaa gatgcttttt 540

ctacaaaccc cgctcttttt accgaaatac aacgggcact aaagcaaccc ctttggaaaa 600

ggttgccaag aatagaggcg gcgcagtaca ttcctttcta tcaacaacaa gattctcata 660

acaagacttt acttaaactt gctaagttag agttcaattt gcttcagtca ttgcacaagg 720

aagagctcag ccatgtgtgc aaatggtgga aagctttcga tatcaagaag aacgcacctt 780

gtttaagaga tagaattgtt gaatgctact tttggggact aggttcaggc tatgagccac 840

agtattcccg ggctagagtt ttcttcacaa aagctgttgc tgttataact cttatagatg 900

acacttatga tgcgtatggt acttatgaag aacttaagat ctttactgaa gctgttgaaa 960

ggtggtcaat tacatgctta gacacacttc cagaatacat gaaaccgata tacaaattat 1020

tcatggatac atacacagaa atggaagaat ttcttgcaaa ggagggaaga acagatctat 1080

ttaactgcgg caaagaattt gtgaaagagt ttgttagaaa cctgatggtt gaagcaaaat 1140

gggcaaatga gggacacata ccaaccactg aagagcatga tccagttgta atcattactg 1200

gcggtgctaa cctgcttaca acaacttgtt atcttggcat gagtgatata ttcacaaaag 1260

agtctgtcga atgggctgtc tctgcacctc ctcttttag atactcaggt atacttggtc 1320

gacgcctaaa tgatctcatg acccacaagg ccgagcaaga aagaaaacat agttcatcga 1380

gccttgaaag ttatatgaag gaatataatg tcaatgagga gtatgcccaa accttgattt 1440

acaaggaagt agaagatgtg tggaaagata taaaccgaga gtacctcaca actaaaaaca 1500

ttccaaggcc gttattgatg gctgtgatct atttgtgcca gtttcttgaa gttcaatatg 1560

caggaaagga taacttcaca cgtatgggag acgaatacaa acatctcata aagtctctac 1620

tcgtttatcc tatgagtata tgaggatcc 1649

<210> 14
<211> 549
<212> PRT
<213> Artificial Sequence

<220>
<223> Deduced amino acid sequence of the amorphadiene synthase encoding gene, between start and stop codon, obtained by PCR with primers C and D

<400> 14

```
Thr Met Ala Leu Thr Glu Glu Lys Pro Ile Arg Pro Ile Ala Asn Phe
                  5                  10                  15
Pro Pro Ser Ile Trp Gly Asp Gln Phe Leu Ile Tyr Gln Lys Gln Val
             20                  25                  30
Glu Gln Gly Val Glu Gln Ile Val Asn Asp Leu Lys Lys Glu Val Arg
         35                  40                  45
Gln Leu Leu Lys Glu Ala Leu Asp Ile Pro Met Lys His Ala Asn Leu
     50                  55                  60
Leu Lys Leu Ile Asp Glu Ile Gln Arg Leu Gly Ile Pro Tyr His Phe
 65                  70                  75                  80
Glu Arg Glu Ile Asp His Ala Leu Gln Cys Ile Tyr Glu Thr Tyr Gly
                 85                  90                  95
Asp Asn Trp Asn Gly Asp Arg Ser Ser Leu Trp Phe Arg Leu Met Arg
            100                 105                 110
Lys Gln Gly Tyr Tyr Val Thr Cys Asp Val Phe Asn Asn Tyr Lys Asp
        115                 120                 125
Lys Asn Gly Ala Phe Lys Gln Ser Leu Ala Asn Asp Val Glu Gly Leu
    130                 135                 140
Leu Glu Leu Tyr Glu Ala Thr Ser Met Arg Val Pro Gly Glu Ile Ile
145                 150                 155                 160
Leu Glu Asp Ala Leu Gly Phe Thr Arg Ser Arg Leu Ser Ile Met Thr
                165                 170                 175
Lys Asp Ala Phe Ser Thr Asn Pro Ala Leu Phe Thr Glu Ile Gln Arg
            180                 185                 190
Ala Leu Lys Gln Pro Leu Trp Lys Arg Leu Pro Arg Ile Glu Ala Ala
        195                 200                 205
Gln Tyr Ile Pro Phe Tyr Gln Gln Gln Asp Ser His Asn Lys Thr Leu
    210                 215                 220
Leu Lys Leu Ala Lys Leu Glu Phe Asn Leu Leu Gln Ser Leu His Lys
225                 230                 235                 240
Glu Glu Leu Ser His Val Cys Lys Trp Trp Lys Ala Phe Asp Ile Lys
                245                 250                 255
Lys Asn Ala Pro Cys Leu Arg Asp Arg Ile Val Glu Cys Tyr Phe Trp
            260                 265                 270
Gly Leu Gly Ser Gly Tyr Glu Pro Gln Tyr Ser Arg Ala Arg Val Phe
        275                 280                 285
Phe Thr Lys Ala Val Ala Val Ile Thr Leu Ile Asp Asp Thr Tyr Asp
    290                 295                 300
Ala Tyr Gly Thr Tyr Glu Glu Leu Lys Ile Phe Thr Glu Ala Val Glu
305                 310                 315                 320
Arg Trp Ser Ile Thr Cys Leu Asp Thr Leu Pro Glu Tyr Met Lys Pro
                325                 330                 335
Ile Tyr Lys Leu Phe Met Asp Thr Tyr Thr Glu Met Glu Glu Phe Leu
            340                 345                 350
Ala Lys Glu Gly Arg Thr Asp Leu Phe Asn Cys Gly Lys Glu Phe Val
        355                 360                 365
Lys Glu Phe Val Arg Asn Leu Met Val Glu Ala Lys Trp Ala Asn Glu
    370                 375                 380
Gly His Ile Pro Thr Thr Glu Glu His Asp Pro Val Val Ile Ile Thr
385                 390                 395                 400
Gly Gly Ala Asn Leu Leu Thr Thr Thr Cys Tyr Leu Gly Met Ser Asp
                405                 410                 415
Ile Phe Thr Lys Glu Ser Val Glu Trp Ala Val Ser Ala Pro Pro Leu
            420                 425                 430
Phe Arg Tyr Ser Gly Ile Leu Gly Arg Arg Leu Asn Asp Leu Met Thr
        435                 440                 445
His Lys Ala Glu Gln Glu Arg Lys His Ser Ser Ser Ser Leu Glu Ser
    450                 455                 460
Tyr Met Lys Glu Tyr Asn Val Asn Glu Glu Tyr Ala Gln Thr Leu Ile
```

```
465                 470                 475                 480
Tyr Lys Glu Val Glu Asp Val Trp Lys Asp Ile Asn Arg Glu Tyr Leu
                485                 490                 495
Thr Thr Lys Asn Ile Pro Arg Pro Leu Leu Met Ala Val Ile Tyr Leu
            500                 505                 510
Cys Gln Phe Leu Glu Val Gln Tyr Ala Gly Lys Asp Asn Phe Thr Arg
        515                 520                 525
Met Gly Asp Glu Tyr Lys His Leu Ile Lys Ser Leu Leu Val Tyr Pro
    530                 535                 540
Met Ser Ile Gly Ser
545
```